HSB Project 2
Benzene ADME Phenotype and Haplotype Association Analyses

Project Leader
John E. French, Ph.D.

Background and Rationale
Benzene metabolites are toxic to the bone marrow of humans and rodents and exposure is associated with hematopoietic diseases (e.g., acute myeloid leukemia and non-Hodgkin's lymphoma). In rodent toxicology and carcinogenesis studies benzene is a multi-site carcinogen by multiple routes of exposure (Huff *et al.*, 1989). The strain specific pattern of ADME kinetics may be determined by individual genetic variation in genes that affect ADME and tissue specific susceptibility to benzene reactive metabolites. Rodent models have been critical to elucidating the potential mechanisms of metabolism and toxicity to benzene. Benzene is an ideal model human toxicant to evaluate because chemical disposition and metabolism have been studied in a number of rodent models and genetically diverse human populations (Bauer *et al.*, 2003; Kim *et al.*, 2007; Lan *et al.*, 2009; Recio *et al.*, 2005; Sabourin *et al.*, 1987).

Key Issues
Conventional ADME studies use 3-5 animals per time point to estimate ADME kinetic parameters by determining the mass balance of the isotope-labeled or non-isotopic labeled test chemical from a single exposure and route at specified intervals. Discrete sampling of individual groups of animals to calculate the time dependent means per group of animals may not have sufficient power for QTL analysis. Alternatively, serial sampling of small volumes of blood over time at different exposure levels below and near saturation of metabolic capacity will provide more variance and improve estimations of bioavailability across strains for correlation with toxicity. These studies and alternate approaches should provide insight into optimizing ADME study designs to aid haplotype association mapping (HAM) and identification of highly penetrant genes that modify these phenotypes.

The phase 1, 2, and 3 genes related to metabolism and transport are considered *a priori* to be related to phenotypic differences between strains. However, we do not know what other genes or regulatory sequences are that may impact ADME phenotypes. Analyses of haplotype-phenotype associations may identify other critical genic or nongenic sequences that modify ADME phenotypes.

The optimal number of strains and approaches for optimizing quantitative measures of ADME phenotypes is also poorly understood. Based upon available resources, we elected to investigate 17 genetically diverse strains along with the B6C3F1/J hybrid strain, which has been extensively investigated. If the 18 strains contained within this cohort are insufficient to obtain sufficient power to identify quantitative trait loci and analyze underlying sequence to identify candidate quantitative genes, we will incorporate more strains into the ADME studies as necessary. Statistical calculations suggest that 30-50 strains may be required to obtain 50-80% power for QTL identification.

Hypothesis

Individual strain genetic variation will result in significant differences in benzene induced ADME kinetic parameters in multiple strains of mice that reflect variation and orthologous genes observed in the human population.

Approach and Specific Aims

Conventional ADME studies measure mass balance from exposure (route specific) to a radioactive or nonradioactive labeled test chemical that requires collection of both blood and tissue samples at specific time intervals. These studies also require multiple animals to be killed at selected intervals for time point specific measurements. Here, we will follow a conventional approach using male and female mice of 18 inbred strains (Frazer *et al.*, 2007; Yang *et al.*, 2007) to determine the potential for identification of genetic sequences associated with the benzene ADME phenotypes. These 18 strains are: C57BL/6J, C3H/HeJ, B6C3F1J, BALB/cByJ, 129S1/SvImJ, A/J, AKR/J, FVB/NJ, BTBRT+FJ, DBA/2J, KK/HiJ, NOD/ShiLtJ, NZW/LacJ, PWD/PhJ, PWK/EiJ, CAST/EiJ, MOLF/EiJ and WSB/EiJ. A haplotype association mapping (HAM) strategy (Harrill *et al.*, 2009; Schadt *et al.*, 2003) will be used to find genetic determinants underlying strain-specific differences in ADME parameters calculated from the quantitative measures at each time interval for males from the 18 strains. Briefly, haplotype associations are calculated using a modified F-statistic based upon strain dependent polymorphisms (SDP)-phenotype pairings at each 3-SNP window across a genome wide SNP dataset with high confidence imputed genotype calls. Association scores (or negative $\log_{10}$ p-values*)* will be plotted across the mouse genome to visualize significance as a function of physical position. Even though this targeted study may be underpowered compared to some HAM studies in the literature (Tsaih and Korstanje, 2009), genomic intervals with association scores greater than 4 will be determined and explored further. Significant associated sequence intervals will be identified using the BioMart feature of Ensembl using NCBI build 37 (http://www.ensembl.org).

Specific Aims

1) Determine the statistical differences between strains for quantitative measures of ADME kinetic parameters.
2) Perform genome wide haplotype association mapping studies to determine potential quantitative trait loci and identify associated candidate sequences for further review and characterization.
3) Analyze candidate genes and flanking regions using bioinformatic approaches to identify regulatory and coding sequences that may have a functional role in expression or function of ADME kinetics.

Significance and Expected Outcome

Analysis of male mice from 18 inbred strains will be used to determine if strain differences exist by comparing area under the curve (AUC) for total [14C] benzene-equivalents, maximum concentration of [14C] radioactivity reached in whole blood following oral administration of [14C] benzene (Cmax), and time to reach Cmax (Tmax) following oral administration of [14C] benzene as described in the previous project summary. Preliminary results show significant strain variation, with apparent Cmax values ranging from 226-1580

nmol-eq/mL (approximately 7-fold) and Tmax values ranging from less than 5 to 17.5 min (3.5-fold). Calculated AUC values ranged from 23 to 239 µmol*min/mL (> 10-fold). Values for the reference strain (C57BL/6J) were 359 nmol-eq/mL at 17 min with an AUC of 45 µmol*min/mL. Pharmacokinetic values represent a mixture of parent benzene and metabolites circulating systemically. Based on the available data generated in the 2-fold difference selection protocol for Cmax, Tmax and AUC values, males from the several strains have been selected for further study. These strains are: NZW/LacJ (high AUC, high Cmax), KK/HiJ (high Cmax), A/J (low Tmax), KK/HiJ (low Tmax), and NOD/ShiLt (low Tmax). Female mouse strains selected for further studies are: PWD/PhJ (low AUC), A/J (low Cmax), DBA/2J (low Cmax), CAST/EiJ (low Cmax), NOD/ShiLtJ (low Cmax). HPLC-radiometric analyses of bladder urine collected after administration of [14C] benzene detected a mixture of metabolites. At least three glucuronide-conjugate metabolites of [14C] benzene have been tentatively identified in the urine of the reference strain at 2 h post dose. Preliminary studies of total [14C] radioactivity disposition in bone marrow from femurs show a similar Tmax for [14C] radioactivity in bone marrow as [14C] radioactivity in systemic blood, albeit at lower levels. These data will be used for haplotype association mapping across each strain and each chromosome for each of the ADME kinetic parameters to identify probable causally related genetic variants.

Current and Future Activities

The kinetic parameter data obtain from the males of the 18 sequenced strains are currently being analyzed by haplotype association analysis to determine if quantitative trait loci can be detected and candidate genetic variants can be identified for functional validation. Similar analyses will be carried out for the females of these strains. Studies with the cardiotoxin, Bis (2-chloromethoxy) methane, which has a metabolite, thiodiglycolic acid, in common with other cardiotoxins, will be investigated in the next series of ADME studies using these genetically diverse mouse strains. If required to complete haplotype association analyses, we will test up 12-20 additional inbred strains for each test agent in a simplified ADME design with serial bleeds of each animal to increase estimates of strain dependent variation.

References

Bauer AK, Faiola B, Abernethy DJ, Marchan R, Pluta LJ, Wong VA *et al* (2003). Genetic susceptibility to benzene-induced toxicity: role of NADPH: quinone oxidoreductase-1. *Cancer Res* **63:** 929-35.

Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ *et al* (2007). A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* **448:** 1050-3.

Harrill AH, Watkins PB, Su S, Ross PK, Harbourt DE, Stylianou IM *et al* (2009). Mouse population-guided resequencing reveals that variants in CD44 contribute to acetaminophen-induced liver injury in humans. *Genome Res* **19:** 1507-15.

Huff JE, Haseman JK, DeMarini DM, Eustis S, Maronpot RR, Peters AC *et al* (1989). Multiple-site carcinogenicity of benzene in Fischer 344 rats and B6C3F1 mice. *Environ Health Perspect* **82:** 125-63.

Kim S, Lan Q, Waidyanatha S, Chanock S, Johnson BA, Vermeulen R *et al* (2007). Genetic polymorphisms and benzene metabolism in humans exposed to a wide range of air concentrations. *Pharmacogenet Genomics* **17:** 789-801.

Lan Q, Zhang L, Shen M, Jo WJ, Vermeulen R, Li G *et al* (2009). Large-scale evaluation of candidate genes identifies associations between DNA repair and genomic maintenance and development of benzene hematotoxicity. *Carcinogenesis* **30:** 50-8.

Recio L, Bauer A, Faiola B (2005). Use of genetically modified mouse models to assess pathways of benzene-induced bone marrow cytotoxicity and genotoxicity. *Chem Biol Interact* **153-154:** 159-64.

Sabourin PJ, Chen BT, Lucier G, Birnbaum LS, Fisher E, Henderson RF (1987). Effect of dose on the absorption and excretion of [14C]benzene administered orally or by inhalation in rats and mice. *Toxicol Appl Pharmacol* **87:** 325-36.

Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V *et al* (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422:** 297-302.

Tsaih SW, Korstanje R (2009). Haplotype association mapping in mice. *Methods Mol Biol* **573:** 213-22.

Yang H, Bell TA, Churchill GA, Pardo-Manuel de Villena F (2007). On the subspecific origin of the laboratory mouse. *Nat Genet* **39:** 1100-7.